COMMENTS ON THE ROLE OF MOLECULAR GENETICS IN POLYMER MATERIALS SCIENCE

MAURILLE J. FOURNIER,[a] THOMAS L. MASON[a] and DAVID A. TIRRELL[b]
Departments of Biochemistry[a] and Polymer Science and Engineering,[b] University of Massachusetts, Amherst, MA 01003

The most fundamental goal of the synthetic chemist is control of molecular architecture. With respect to small molecules (i.e., those of molecular weight less than a few thousand), this means absolute control of chemical connectivity and stereochemistry – complete specification of molecular structure. But in macromolecular chemistry, controlled architecture has meant something quite different. Because polymerizations are in general statistical processes, conventional polymeric materials are characterized by substantial heterogeneity in chain length, sequence and stereochemistry [1]. Control is exercised in a statistical sense only, and considerable skill is required to control even the average properties of the chain population and the dispersity in those properties.

The most powerful of the traditional approaches to the control of chain architecture are living polymerization [2] and Ziegler-Natta polymerization [3]. Each constitutes a partial solution to the architectural problem. Living polymerizations afford control of average molecular weights and the breadth of the molecular weight distribution, and Ziegler-Natta catalysis provides an effective means of controlling the stereochemistry of the chain. Historically, technological exploitation of each of these fundamental advances proceeded rapidly, and each discovery led quickly to the introduction of important new products. Ziegler-Natta polymerization is now used to produce crystalline polyolefins in quantities that exceed 25 billion pounds per year, and living polymerization provides the basis for the production of several important classes of thermoplastic elastomers. There is cause for optimism that further synthetic advances, and specifically those that lead to increased control of chain architecture, will in turn create significant new technological opportunities.

Recent developments in the synthesis, cloning and expression of artificial genes offer the prospect of just such an advance. It appears likely that genetic strategies will allow the preparation of virtually any amino acid copolymer, regardless of its length, composition or sequence. In principle, then, the synthetic problem can be solved completely for this class of polymers. On the other hand, implementation of this strategy in the preparation of specific target polypeptides remains far from trivial, and reflects our incomplete understanding of the factors that control the stability and expression of artificial genes and the accumulation of heterologous proteins in cellular hosts. The discussion that follows provides first a brief overview of the problems that must be overcome if genetic strategies are to earn an important role in polymer materials science, and then a few speculative remarks on what that role might be.

POTENTIAL PROBLEMS

Figure 1 outlines the important steps in protein biosynthesis, and serves as a basis for discussion of potential problems that may arise in attempts to exploit this process to prepare new polymeric materials. Those problems are:

Mat. Res. Soc. Symp. Proc. Vol. 218. ©1991 Materials Research Society

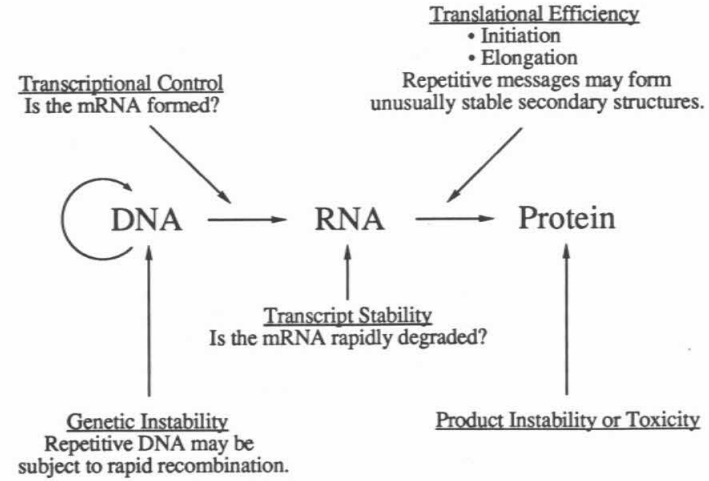

Figure 1. Key steps in protein biosynthesis, highlighting problem areas in the exploitation of this process to produce new polymeric materials.

Genetic Instability

The replication of DNA represents a substantial metabolic burden on the cell. Introduction (and replication) of foreign DNA increases that burden, most often without conferring any competitive growth advantage on the host. Host cells that do not maintain the foreign sequence (or a portion of it) may then grow and divide more rapidly than their siblings, resulting in total or partial loss of the artificial coding sequence from the cell population. This problem appears likely to be particularly acute for the repetitive coding sequences of special interest in materials science, since such sequences may be especially prone to rapid rearrangement via homologous recombination. Several tandemly repeated sequences have indeed been found to be unstable in bacterial hosts [4].

RNA Synthesis and Stability

Efficient protein synthesis requires an adequate steady-state concentration of the messenger RNA transcript of the coding sequence. The rate of mRNA synthesis is controlled in large measure by the promoter chosen to direct transcription, but it appears that the rates of RNA synthesis and degradation are sequence-dependent as well [5].

Translational Efficiency

Translation of the message into the protein product of interest involves discrete initiation, elongation and termination steps. Initiation of translation in bacterial hosts requires a "start" codon, usually AUG, located 4-15 nucleotides downstream of a ribosomal binding site that positions the message properly on the small ribosomal subunit through base-pairing interactions with the 3'-region of the 16S ribosomal RNA. On this point there is general agreement. What is not so clear is whether or not there are further sequence requirements for initiation. Recent work from several laboratories has suggested that there are strong compositional preferences in natural *E. coli* startsites, both 5' and 3' to the AUG [6], but a persuasive mechanistic rationale for such preferences has not been described.

Our understanding of the connection between mRNA sequence and rate of protein elongation is similarly incomplete. Experimental results that speak directly to this point are rare [7]. In principle, sequence sensitivity in the elongation rate might arise from sequence-dependent secondary structure in the message, or from variations in the availability of the cognate transfer RNAs. Although clear correlations remain to be demonstrated, a prudent approach to gene design incorporates both considerations into the process of codon selection.

Protein Toxicity or Instability

Even those genetic strategies that succeed at the nucleic acid level can fail in the final analysis if the protein product is excessively toxic or unstable. Given the abiological nature of many of the protein sequences of materials interest, toxicity is likely to be particularly hard to predict. And while systematic investigations of the relations between protein sequence and turnover rate have been undertaken [8], it seems certain that general, reliable predictive schemes will be slow to emerge.

IMPLICATIONS FOR MATERIALS RESEARCH AND DEVELOPMENT

The foregoing discussion summarizes the problems that can frustrate the routine use of genetic strategies to prepare new polymeric materials. Solutions to these problems can be found, however, as shown by our own work [9] and that of others [10]. At present, these solutions are developed largely on an *ad hoc* basis, and the field will be well served by research programs that address not only specific materials objectives but also the fundamental biological issues discussed above. There is reason for optimism that such efforts will in the not-too-distant future lead to a set of reliable rules for the design and expression of important new protein-based polymeric materials.

What might the impact of these new materials be? Any answer to such a question should take account of three characteristic features of polymers produced in this way: i). their structural homogeneity, ii). their chemical functionality, and iii). their biological origins.

Their structural homogeneity makes these new materials powerful tools for use in fundamental studies of macromolecular behavior. Faced with the limitations of conventional synthetic methods, polymer materials scientists have been forced to deal with heterogeneous mixtures of chains and to interpret their observations not only in terms of the average structure of the chain population, but also in terms of any dispersity in chain length, composition, sequence and stereochemistry. The decoupling of these structural variables can be difficult or impossible. Genetic strategies on the other hand should allow the preparation of homogeneous populations of chains designed precisely to test fundamental ideas about the relations between chemical structure and materials properties. Crystallization, diffusion, adsorption and phase behavior seem to be particularly attractive areas for investigation. Much of our own work has been, and will continue to be, directed toward these objectives.

In comparison with more traditional approaches to the control of macromolecular architecture, genetic strategies offer enormous advantages in terms of the chemical functionality that can be accommodated. This leads at once to important possibilities for the chemical modification of appropriately designed protein chains to produce new materials with interesting optical, electronic, or liquid crystalline properties. We anticipate that many of the new materials that will emerge from this approach will be highly modified proteins, in which the polypeptide backbone serves largely to organize the solid, and to create appropriate spatial arrangements of the functional groups of interest.

Finally, the biological origins – and the potential biological functions – of protein-based polymers should not be overlooked. Creative use of genetic strategies should lead not only to biodegradable polymers, but also to many interesting and important hybrid proteins, in which

structural and functional domains are combined. Important progress along these lines has been made by Cappello and coworkers, in the construction of silk-like proteins that incorporate the cell-binding domain from the extracellular matrix protein fibronectin [11]. These hybrid proteins can be used to coat inert surfaces, and after coating, will promote cell binding with an efficiency comparable to that of natural fibronectin. The immediate application of such materials is likely to be in cell culture, but the broader implications of the approach extend to many areas of biomaterials design.

REFERENCES

1. G. Odian, Principles of Polymerization, 2nd ed., Wiley, New York, 1981.
2. M. Szwarc, *Adv. Polym. Sci.* **49**, 1 (1983).
3. J. Boor, Jr., Ziegler-Natta Catalysts and Polymerizations, Academic Press, New York, 1979.
4. S. C. Gupta, H. L. Weigh and R. L. Somerville, *Biotechnology* **9**, 602 (1983).
5. For a brief review, see G. Brawerman, *Cell* **57**, 9 (1989).
6. M. Dreyfus, *J. Mol. Biol.* **204**, 79 (1988) and references therein.
7. G. A. Gutman and G. W. Hatfield, *Proc. Natl. Acad. Sci., USA* **86**, 3699 (1989) and references therein.
8. A. Varshavsky, A. Bachmair, D. Finley, D. Gonda and I. Wünning, in Ubiquitin, M. Rechsteiner, ed., Plenum, New York, 1988, p. 287; S. A. Goff, R. Voellmy and A. L. Goldberg, ibid. p. 207.
9. K. P. McGrath, D. A. Tirrell, M. Kawai, T. L. Mason and M. J. Fournier, *Biotech. Prog.*, **6**, 188 (1990); K. P. McGrath, M. J. Fournier, T. L. Mason and D. A. Tirrell, *Polym. Prepr.* **30(1)**, 190 (1990); H.S. Creel, M.J. Fournier, T.L. Mason and D.A. Tirrell, *Macromolecules*, in press.
10. Preprints of papers, Symposium on Protein-Based Polymers, American Chemical Society, Boston, 1990; *Polym. Prepr.* **30(1)**, 176 ff (1990); J. Cappello, J. Crissman, M. Dorman, M. Mikolajczak, G. Textor, M. Marquet and F. Ferrari, *Biotech. Prog.* **6**, 198 (1990); I. Goldberg, A.J. Salerno, T. Patterson and J.I. Williams, *Gene* **80**, 305 (1989).
11. J. Cappello and J. W. Crissman, *Polym. Prepr.* **30(1)**, 193 (1990).